

Draft. Please do not quote without permission. Comments welcome.

Psychophysical Methods and the Evasion of Introspection

M. Chirimuuta
Dept. History & Philosophy of Science
1017 Cathedral of Learning
4200 Fifth Avenue
University of Pittsburgh, Pittsburgh, PA 15260

mac289@pitt.edu

Abstract

While introspective methods went out of favour with the decline of Titchener's analytic school, many important questions concern the rehabilitation of introspection in contemporary psychology. Hatfield (2005) rightly points out that introspective methods should not be confused with analytic ones, and goes on to describe their "ineliminable role" in perceptual psychology. Here I argue that certain methodological conventions within psychophysics reflect a continued uncertainty over appropriate use of subjects' perceptual observations and the reliability of their introspective judgements.

My first claim is that different psychophysical methods do not rely equally on the introspective capabilities of experimental subjects. I contrast "minimally-introspective" tasks with "introspection-heavy" ones. It is only in the latter, I argue, that introspection can be said to have a non-trivial role in the subjects' performance. My second claim is that my rough-and-ready distinction maps onto a number of important "dichotomies" in vision science (Kingdom and Prins 2009). Not coincidentally, the introspection-heavy categorisation captures many of the tasks typically considered less able to yield useful information regarding the processes underlying visual sensation.

1. Introduction

Recent work on introspection in psychology has been careful to separate the specific commitments of Titchener's analytical school from the discussion of introspection more generally. For example, Hatfield (2005) defines introspection broadly as, "a mental state or activity in or through which persons are aware of properties or aspects of their own conscious experience" (p.260). He later defines introspection as, "deliberate and immediate attention to certain aspects of phenomenal experience," arguing that, "it continues to be used as a source of evidence in perceptual and cognitive psychology" (p.279). In this paper I will challenge the appropriateness of Hatfield's definitions in the branch of perceptual psychology known as psychophysics¹, offering an alternative account.

¹ Psychophysics is defined by Gescheider (1997) as "the scientific study of the relation between stimulus and sensation." The disciplinary demarcation between psychophysics and perceptual

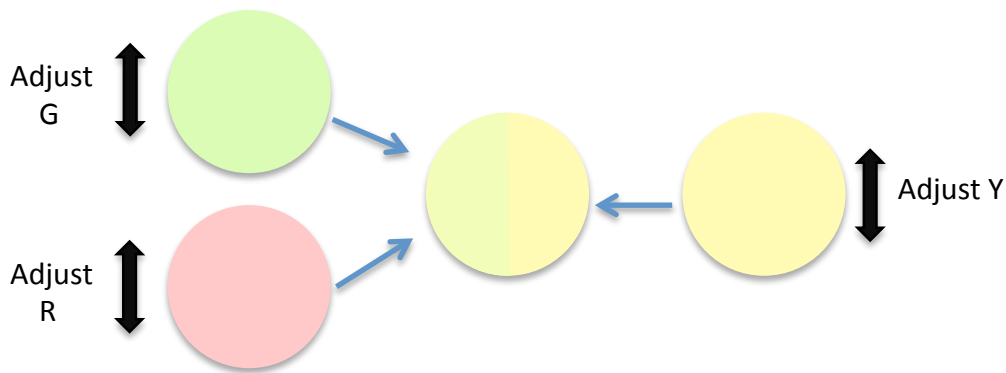


Figure 1

A Metameric Match Experiment. The subject is asked to adjust the intensities of R, G and Y monochromatic lights so that the yellows are indistinguishable.

Hatfield discusses the psychophysical task of metamerism colour matching (see fig. 1) as one example of a perceptual experiment reliant on introspection (p.278). However, it is reasonable to question whether Hatfield's definitions can effectively target those activities which are introspective, or if they are too permissive and encompass a range of activities ordinarily thought of as just perceptual and not reliant on introspection. First, though, there is an important exegetical question over how to understand Hatfield's claim that introspection is a source of evidence in experiments such as the metamerism matching one.

One possible reading, which I reject, is that Hatfield just points to the fact that psychophysics, unlike behaviourist psychology assumes and moreover requires that experimental subjects have conscious perceptual experiences². For the mission of psychophysics, an experimental approach to the mind inaugurated by Fechner (1860), is to chart and measure the physical energies needed to elicit specific conscious perceptual states. But I strongly doubt that Hatfield intends to characterise methods in psychology as introspective purely in terms of a contrast with the behaviourist research program. In fact, Hatfield is in agreement with Danziger (1980) that our understanding of introspection in psychology has been distorted by the behaviourist reaction to Titchener's analytical school.

What is more, the mere having of conscious states is a different thing from the possession of some ability to report reliably on the nature of those states. It is the latter capacity that is typically identified with introspection. For example,

psychology more generally has become somewhat blurry in recent years, with many experiments that are classified as psychophysical dealing with complex perceptual states, not just simple sensation.

² Not to say that the behaviourist psychologists all assumed that human beings were unconscious zombies, but that their experimental methods were indifferent to the presence or absence of consciousness.

Schwitzgebel's (2011) sceptical case against introspection does not target the idea that we have conscious states, or that those states are important to our mental economy, but is concerned to argue that to a greater extent than we care to admit the contents of those states are indeterminate or unknown to us. My interpretation of Hatfield's notion of introspection will hinge on this point. I understand his claim that contemporary perceptual psychology relies on introspective evidence to be the claim that psychologists exploit their subjects' introspective ability in order to glean information about the human perceptual system, and furthermore it is a presupposition of this experimental practice that subjects are competent introspectors in the sense that they are capable of giving verbal or motor responses which reliably indicate the presence or absence of particular features of their conscious experiences. For example, a psychophysical experiment which measures the absolute detection threshold for a dim spot of light is said to be reliant on the subject's capacity to introspect in the sense that her subjective awareness of the spot is a crucial data point that the experimenter has access to because of the subject's capacity to introspect. And thus the experimenter must assume that the subject can faithfully indicate those times that the spot enters her conscious field of view.

Yet a problem with this account is that it is not clear how it can be employed to distinguish introspection from ordinary perception, for doesn't the subject's activity in the detection task just boil down to her *looking* for a dim spot of light? This worry could prompt us to take Hatfield as endorsing a more restrictive definition. For Hatfield also suggests that what characterises introspection over ordinary perception is that one attends to one's experience of an object, not just the object itself (p. 279, "immediate attention to... phenomenal experience"). This makes introspection importantly different from perception because as many would have it, perception is generally "transparent" and our perceptual encounter with the world is not interrupted with moments of attention to experience itself. The difficulty with this reading is that it then becomes unclear how the more restrictive definition of introspection could apply to many of the psychophysical tasks that Hatfield wants it to apply to, such as stimulus detection and the metamerism matching experiment. Subjects perform such tasks by directing their attention to external stimuli, namely the coloured lights, and need not attend to their own phenomenal experience, *qua* experience. Nor do they need to consider their experience in a more fine grained or detailed way than in ordinary perception.

In short, the problem is that while Hatfield's restrictive definition has the virtue allowing one to demarcate introspection from perception, it cannot reasonably be applied to the range of psychophysical tasks that Hatfield claims it does. And furthermore a case could be made that it should not apply to any perceptual experiment, since these generally involve attention to external objects, not attention to phenomenal experience itself. Yet the more liberal definition makes all perceptual activity concurrently introspective in a somewhat trivial sense.

It strikes me that a different approach to defining introspection -- in the context of psychophysics -- is needed, one that does not characterise introspection in terms of an object of attention or focus of awareness. In this paper I propose that the tasks that should be said to involve introspection are the *ones which rely on experimental subjects' capacity to analyse and compare sensory experiences that bear non-obvious relationships of similarity and difference to each other*. Thus on my account introspection can be part of the process of perceiving and attending to an external object, and need not be overtly directed at phenomenal experience. The subject may interpret her task to be simply that of attending to the external stimulus, but she can be reporting on aspects of her phenomenal experience nonetheless. It is also a feature of my view that the extent to which tasks rely on introspection is a matter of degree. In the next section I give a set of examples of common psychophysical tasks that are either "introspection-heavy" or "minimally-introspective". In the third section I describe how the cluster of introspection-heavy tasks – though not described in this way by scientists themselves – has commonly attracted suspicion from psychophysicists as being less likely to produce data that is "objective" and informative about neural mechanisms. I ask whether this is mere coincidence, or if the methodological norms of psychophysics reflect a certain wariness towards introspection.

2. Introspection in Psychophysics as Controlled Comparison

2.1 Examples Of Introspectively Demanding and Undemanding Psychophysical Tasks.

The *metameric match paradigm*, illustrated in figure 1 has been used to diagnose specific types of colour vision deficiency since the late 19th century. Differences in the number of retinal cone types an individual has, and the spectral sensitivities of those cone classes, lead to measurable differences in the proportion of red to green in a composite light that he or she judges to look identical to a yellow monochromatic standard. Note that in this task the only perceptual judgment that the subject need make is over whether the composite and monochromatic light are visually indistinguishable. If the lights are presented as abutting (as in fig. 1) then the subject simply has to judge whether or not the colour field is homogeneous. No attention to the specific qualities of the perceived colour is required.

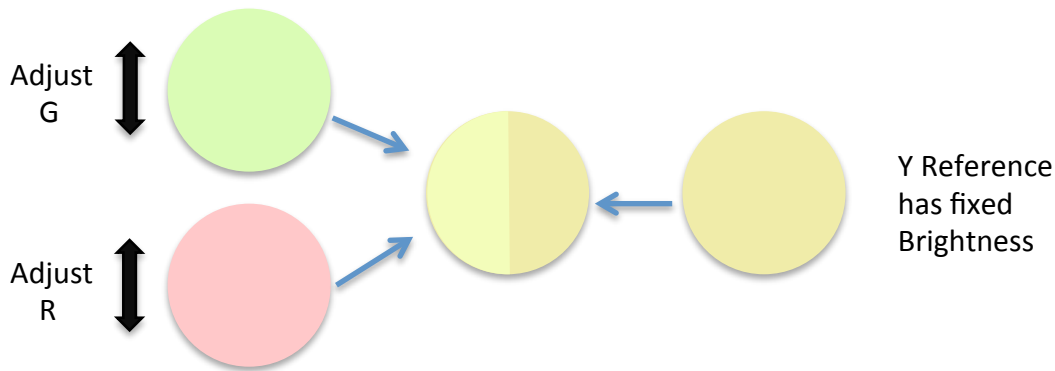


Figure 2

An Asymmetric Match Experiment. The subject is asked to adjust the proportions of R and G monochromatic lights so that the yellows match in hue. The intensity of the Y light is fixed, so the yellows cannot be matched for brightness and are therefore distinguishable even when hue is judged to be equivalent.

Contrast this task with an *asymmetric match paradigm* (fig. 2). In this case the two central stimuli are not matched for luminosity but the subject must say whether or not they match in hue regardless of their visible difference in brightness. This requires that the subject analyse her experience of the two colours in terms of separate dimensions of hue and brightness, and make a judgment as to the identity of just one of these dimensions, disregarding the difference in the other. Thus the subject must make a series of comparisons between pairs of stimuli in order to find the pair that holds the unique but non-obvious relationship of sameness of hue. This relationship is non-obvious in that it is not marked by a simple defining characteristic like a homogenous spatial profile.

It should be fairly intuitive that this task is “introspection-heavy” in a way that the metameric matching task is not. The contrast between these two tasks points us to the central intuition behind my new characterisation of introspection. The idea is that the metameric matching task is “minimally-introspective” because it can be performed without any careful comparison of the phenomenal qualities one experiences on presentation of the two stimuli. The metameric paradigm relies on introspection only in the minimal sense that it assumes the subject can know and reliably report when her conscious visual field is homogeneous with respect to colour³. The asymmetric matching task, on the other hand, is “introspection-heavy” because it does require this careful comparison of sensory experiences that bear non-obvious relationships of similarity and difference to each other.

Asymmetric matching paradigms have been used to study achromatic perception of lightness and darkness (fig. 3a, see Gilchrist 2004) and to study colour constancy. Figure 3b gives an example of an asymmetric task in which the observer views a

³ i.e. relies on introspection defined in the first, permissive sense. To reiterate the discussion of section 1, the problem with the minimal notion of introspection is that it cannot distinguish introspection from ordinary perception.

scene under two different lighting conditions. She is instructed to adjust the colour of the central patch in one image until it looks as if made from the same paper as the central patch in the other (Foster, 2011). Importantly, even when the patches are matched there will still be a visible difference in colour between them, and the experiment relies on the subject having a clear sense of what sameness of material would look like in spite of these differences. Again, the task is “introspection-heavy” in comparison to a task in which the subject just has to report on the absolute identity or distinguishability of two stimuli. In particular, it relies on the subject’s ability to make a “judgment call” on the one best match, given a range of close contenders which vary along a number of different dimensions. I describe the introspection-heavy tasks as requiring *controlled* comparison because the demand placed on the subject is to perform some kind of analysis and comparison, but within parameters that are pre-specified by the experimenter.

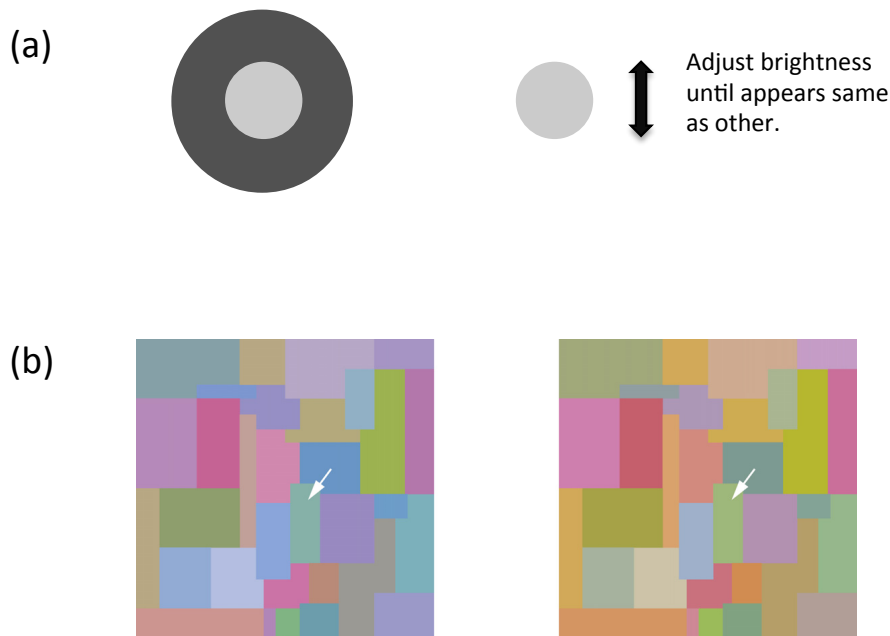


Figure 3

(a) *Achromatic asymmetric match experiment where black annulus influences perceived brightness of one of the circles. Subject is asked to determine point of subjective equality of the brightness of the two circles.*

(b) *Asymmetric colour constancy experiment. Subject is asked to adjust the colour of one of the patches (marked with arrow) until it looks as if it is made from the same paper as the other. (From Foster 2011, permission needed.)*

Another kind of paradigm that intuitively fits the idea of controlled comparison is a rating scale task. In a series of experiments published recently (To et al 2008, 2010, Tolhurst et al 2010) subjects were presented with nearly 300 pairs of photographs – an original and a modified version – and were asked to rate how similar the pairs were on a scale from 0 (completely identical) to any arbitrarily high value (see fig.

4a). In one of these publications, Tolhurst and colleagues (2010) also present results of a simple two-alternative-forced-choice (2-AFC) contrast discrimination experiment in which subjects just had to report which of a pair of otherwise identical photographs contained a small, high contrast central patch (see fig. 4b). They then apply their model of contrast discrimination to the rating scale data. The rating scale task falls under my introspection-heavy category, while the contrast discrimination task is minimally-introspective. In the former, the subject must make a judgement as to the relative similarity of a large number of pairs of stimuli, that differ in different ways, whereas in the latter task she detects the presence or absence of a high-contrast patch in a rather automatic fashion. Figure 4 illustrates how similar stimuli can be used in these two very different experiments, so it is not complexity of stimulus *per se* that determines how introspectively demanding the task is. Rather, the determining factor is the nature of the response that the subject must make to the stimulus. That is, whether the response is simply choice between saying the high contrast patch appeared first or second out of two stimuli, or if it calls for a more careful examination of the perceived properties of the stimuli.

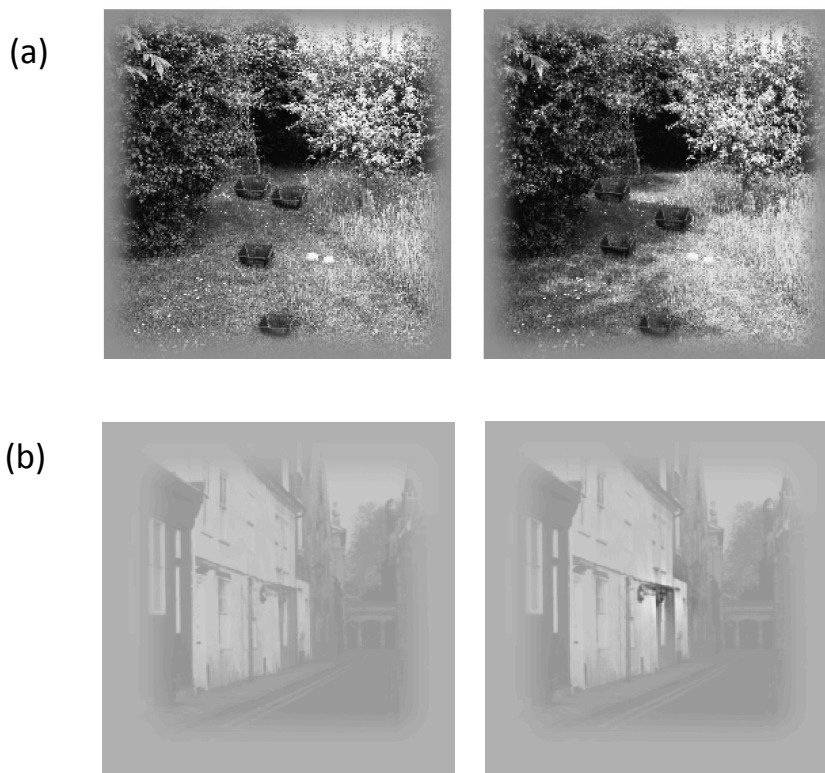


Figure 4

Example of stimuli used by Tolhurst et al. 2010 (a) Rating scale task. For each of 294 image pairs, subjects were asked to rate how similar or different they appeared on a numerical scale of their own devising. (permission needed)

(b) 2-AFC contrast discrimination. Subjects had to report if the high contrast central patch appeared in the first or second stimulus.

Before moving on, I would like to emphasise that my two categories are intended to reflect a qualitative difference in how introspectively demanding these tasks are, and that I will say nothing in this paper about how to quantify this difference, and how it is that introspective demands admit of degree. For example, the question of whether or not metameric matching is even less introspectively demanding than contrast discrimination will be left unanswered. It seems plausible that introspective demands, like attentional demands, come in degrees but I offer no suggestions of how one might measure this. It is also plausible that there will be some tasks that occupy middle ground between my categories and are hard to classify either way. I will not deal with such cases here. My aim in presenting a set of tasks that are intuitively more reliant on introspection than the others has been to highlight one way that introspection may be said to play a role in perceptual psychology, and to this end I have focused on the most clear cut cases.

2.2 Other Classifications of Psychophysical Tasks

One of the attractive things about psychophysics as a subject for philosophy of science is the fact that throughout its short history methodological questions about the best way to measure sensory responses have been debated in a perspicacious way by leading protagonists. Moreover, such controversies still resonate in the living memory of the discipline, and are recounted even in the most recent textbooks. One way in which methodological debates commonly unfold is with a distinction first being drawn between two broad classes of psychophysical techniques, and the relative merits of the two classes are subsequently discussed.

In their textbook Kingdom and Prins (2009) devote a chapter to the “dichotomies” that have been most significant to psychophysicists past and present. The first of these, Brindley’s (1960, 1970) distinction between Class A and Class B observations is particularly relevant to my account of introspection. Brindley characterised Class A observations as any tasks in which the observer just had to report on the absolute similarity or dissimilarity in the appearance of a pair of stimuli. For example, the measurement of the detection threshold for a spot of light is Class A because the subject need only indicate whether the trial in which the spot is present is distinguishable or not from the reference stimulus in which the spot is absent. Likewise, the measurement of the discrimination threshold for the brightness of the spot is also Class A, as it just requires the subject to report if the trial in which the luminosity of the spot is increased looks different from the trial in which the luminosity remained at baseline. In contrast, Brindley (1970:133) categorised as Class B, “[a]ny observation that cannot be expressed as the identity or non-identity of two sensations...”; for example, “all those [observations] in which the subject must describe the quality or intensity of his sensations, or abstract from two different sensations some aspect in which they are alike.”

Brindley's description of Class B observations is interchangeable with my characterisation of introspection-heavy tasks. Indeed, the tasks which I presented as examples of my minimally-introspective category – metamerism matching and contrast discrimination – are Class A, whereas all kinds of asymmetric matching and rating scale tasks are Class B. In essence, both of these categorisation schemes can be understood as drawing a distinction between tasks in which the experimental subject is treated somewhat like a thoughtless measuring instrument, and methods that rely on the subject's status as a critical being who can attend to and reflect on her own conscious states. The point is not that the A/minimally-introspective Class treats the subject as if unconscious, or that it requires the subject to have sensory capacities but not cognitive ones. Rather, it is that the A/minimally-introspective Class makes no demands on any capacity for reflection on and comparison of occurrent sensory states, whereas tasks in the B/introspection-heavy Class do⁴.

To illustrate this, imagine a machine that can read off the conscious sensory states of a subject performing a contrast discrimination task. In order to predict the subject's responses to any trial, all the machine must do is to assign a number to the intensities of the subject's experience of contrast for the central regions of the two different stimuli. If they have the same values the machine predicts the answer is 'same', and if they differ the machine predicts an answer of 'different'. Once the non-trivial problem of reading off individuals' phenomenal states is solved, the rest is uncomplicated! If an equivalent machine were to be built for the rating scale task, the blueprint could not be so simple. There is no one quality of the subject's conscious experience of the stimuli that the machine could measure and use to predict the response. Rather, the machine would have to rely on some complicated model of how various differences in the experienced qualities of the images are weighted against each other to give an impression of greater or lesser degrees of similarity⁵. In other words, a model of introspective comparison and not a simple measurement algorithm.

The mind-reading machine thought experiment again confronts us with the fact that the distinction being drawn is not between tasks that are in no way introspective and those that completely are. Rather, it is about the extent to which these tasks call upon some putative introspective capacity. For the first machine, dealing with the

⁴ In support of this idea that the key distinction in play here is between subject-as-measuring-instrument and subject-as-reflective-being, it is worth noting that Brindley's one example of a psychophysical document explicitly hostile to Class B observations is the 1943 Optical Society of America (OSA) report that, as Stevens (1951:31) relates, "reduces psychophysics to the employment of a human observer as a null instrument under a set of strictly specified conditions" And Brindley's one example of a psychophysicist liberal with regards to Class B is Stevens (1951), who explicitly rejects the OSA definition as too narrow and restrictive (and cf. Helson 1949).

⁵ Interestingly, however, Tolhurst et al. (2010) can predict trends in the similarity rating data to a fair degree of accuracy with a model of entirely unconscious neuronal response functions. The fact that there is "machine" that can predict responses to the contrast discrimination and rating scale experiments, without peering into the conscious states of subjects should not detract from the fact that any hypothetical machine attempting to examine conscious states in order to predict responses would have to treat the two experiments differently.

contrast discrimination experiment, can still peer into the conscious states of observers and this captures some minimal notion of introspective activity. Yet the second machine, dealing with the rating scale experiment, needs not only to determine what the subject experiences, but also to determine what the subject makes of her experience, what is more and less salient about the different qualities presented in her visual phenomenology. This is an introspective undertaking of a weightier kind.

It is hard to say how influential Brindley's distinction has been. It came under immediate criticism from Boynton and Onley (1962) but was clearly accepted in some form by Marks (1978) and Teller (1984), and is discussed at length in Gescheider's (1997) psychophysics textbooks. Kingdom and Prins (2009, p.18) choose not to employ it as an overarching basis for classifying psychophysical experiments because of the problem that certain tasks cannot be classified as either A or B.

Kingdom and Prins' preferred distinction is between tasks that measure performance and those that measure appearance, which they characterise in the following way:

"If the measurement can be meaningfully considered to be better under one condition than under another, then it is a performance measure, if not it is an appearance measure." (p.22)

Performance tasks are any ones designed to chart perceptual "limits" (e.g. contrast discrimination, detection of a spot of light against a differently coloured background). An example of an appearance task is an experiment comparing the strength of the Müller-Lyer illusion with fin angles of 45° and 60°. Even if the length of the central bars appears to be more different when the fins are 45°, there is no sense in which the subject is "better" at the task in that condition. So this Class B observation can also be said to be an appearance measure. Thus there is an overlap with my distinction: appearance tasks tend to be introspection-heavy, and performance tasks tend to be minimally-introspective. But it is not as well matched as is the case with Class A vs. B. In particular, the metameric match task that I classify as minimally-introspective turns out to be an appearance measure.⁶

3. Not All Psychophysical Methods Were Created Equal

All I have argued so far is that there is an intuitive way of differentiating psychophysical tasks that are more reliant on introspection from those that are not, and that my categorisations turn out to be roughly co-extensional with categorisations of tasks developed within the psychophysical tradition. The

⁶ A related dichotomy is Sperling's (et al. 1990) Type 1 vs. Type 2 distinction. In Type 1 experiments the subject's response maybe either correct or incorrect with respect to some physical dimension of the stimulus (e.g. for either is more oblique than line 2). For Type 2 the experimenter is cannot classify responses as correct or incorrect. Note again that the metameric match turns out to be Type 2, even though Class A/minimally-introspective.

question now is what to make of this finding. Is it just a coincidence that the distinctions coincide? It should come as no surprise to the reader that my next point will be that the categories that line up on the introspection-heavy side have tended to meet with more diffidence and suspicion from psychophysicists than those on the other side.

Brindley (1970) presents type A observations as especially informative about the physiological mechanisms underlying perception because they can be used to test “psycho-physical linking hypotheses”, that two stimuli (e.g. yellow monochromatic light, and a certain mixture of red and green lights) will produce the same neural activity and hence the same sensation.

On the relative status of the two classes he writes that,

“The use of Class A observations as a basis for analysing the function of the eye and visual pathway is not controversial; every writer on vision admits, at least by implication, that they can be legitimately used. On the use of the kinds of observation here called Class B, there have been differences of opinion ... The conservative opinion, in its most extreme form, is that only Class A observations are of any value, and in a discussion of visual mechanisms all Class B observations may be entirely disregarded.” (1970, p. 134)

Brindley himself takes this view to be too narrow, but is critical of Stevens’ (1951) “extreme liberal opinion” for failing to make the distinction. Later in the book, when discussing Hering’s opponent theory of colour he writes as if it is still moot whether the kinds of phenomenological reports presented by Hering in support of his theory can actually be taken as evidence for a kind of colour mechanism (p.208).

One might think that this is all besides the point in a discussion about introspection because the reason why the value of Class B observations was held in question was not because they are introspection-heavy, but because their failure to underwrite psychophysical bridge principles. But I do not think that this problem is so disconnected to from the issue of introspection. For if Class B tasks were to be granted some supporting assumptions, like the ones offered for Class A, then one could equally say that they are informative of underlying neural mechanisms. For example, in the case of the asymmetric hue matching experiment, why not assume that when the hue sensation for each stimulus is equal, that is evidence that there is a neural pathway somewhere between the photoreceptors and the cortex that conveys the same message in both cases? This would be a special case of the assumption made in support of inferences from Class A observations that, “whenever two stimuli cause physically indistinguishable signals to be sent from the sense organs to the brain, the sensations produced by these stimuli, as reported by the subject in words, symbols or actions, must also be indistinguishable” (Brindley 1970, p.133).

Yet, Class B observations *are* treated differently. The reason for this difference is likely because Brindley and other theorists (e.g Marks 1978) have been wary of attributing to subjects the kind of introspective powers that would be needed to

analyse hue separately from all other sensory qualities, and determine exactly the point of equivalence of hue. In other words, if these theorists had shared Titchener's faith in the analytical acumen of introspection, they would have had no reason to treat Class B observations differently from Class A.

This pattern of unequal treatment can be seen not just in the discussion of Class A and B observations, but also with respect to the other dichotomies discussed by Kingdom and Prins. They note that it is fairly common for psychophysicists to refer to some tasks as more "objective" or "subjective" than others, with all the value-laden connotations of these terms. Kingdom and Prins explain this usage in the following way:

"All psychophysical experiments are in a trivial sense subjective, because they measure what is going on inside the head, and if this is the intended meaning of the term then the distinction is redundant⁷. The dichotomy is more often invoked, however, to differentiate between different types of psychophysical procedure. The distinction has been used variously to characterize Class A versus Class B observations, tasks for which there is versus tasks for which there is not a correct and an incorrect response⁸, forced-choice versus non-forced-choice procedures, and criterion-dependent versus criterion-free procedures." (p.18-19)

The notion of subjectivity at play here is encapsulated in the idea that experiments are subjective if they are introspection-heavy. For all the tasks on the wrong side of the subjective-objective tracks are ones which rely on the subject's judgments concerning the appearance of the stimuli, involving complex comparisons which cannot be independently verified by examining the physical properties of the stimuli themselves.

To conclude, there is a sense in which the title of this paper is misleading. I have not showed that the psychophysicists have avoided using experimental methods more reliant on introspection, or that the use of such methods has always been questioned. Indeed, when Kingdom and Prins write that, "Both performance-based and appearance-based experiments are important to our understanding of vision. Measures from both types of experiments are probably necessary to fully characterize the system" (p.26), they are articulating a methodological pluralism that many psychophysicists would endorse. However, the crucial point is that the methods on the wrong side of the divide, those more reliant on introspection, continue to need their advocates, whereas those on the other have been accepted without question. This is an indication of the contested status of introspection within the psychophysics tradition.

⁷ Cf. the worry discussed above that all psychophysical experiments rely on introspection in a trivial or "minimal" way, hence the distinction between introspection and perception is made redundant.

⁸ I.e. performance vs. appearance or Sperling's Type 1 vs. Type 2.

References

- Boynton, R. M. and J. W. Onley (1962). "A critique of the special status assigned by Brindley to 'Psychophysical Linking Hypotheses' of 'Class A'." *Vision Research* 2: 383-390.
- Brindley, G. S. (1960, 2nd edition 1970). *Physiology of the Retina and the Visual Pathway*. London, Edward Arnold.
- Danziger, K. (1980). "The History of Introspection Reconsidered." *Journal for the History of the Behavioural Sciences* 16: 241-262.
- Fechner, G. (1860/1966). *Elements of Psychophysics* Holt, Rinehard and Winston.
- Foster, D. H. (2011). "Color Constancy." *Vision Research* 51: 674-700.
- Gescheider, G. A. (1997). *Psychophysics: The fundamentals*. Mahwah NJ, Lawrence Erlbaum.
- Gilchrist, A. (2004). *Seeing Black and White*. Oxford, Oxford University Press.
- Hatfield, G. (2005). *Introspective Evidence in Psychology. Scientific Evidence: Philosophical theories and applications*. P. Achinstein. Baltimore, Johns Hopkins University Press.
- Helson, H. (1949). "Review of 'Introduction to Color'." *Psychological Bulletin* 46(2): 166-169.
- Kingdom, F. A. A. and N. Prins (2009). *Psychophysics: A practical introduction*. Amsterdam, Elsevier Academic Press.
- Marks, L. E. (1978). *The Unity of the Senses*. New York, Academic Press
- Schwitzgebel, E. (2011). *Perplexities of Consciousness*. Cambridge MA, MIT Press.
- Sperling, G., B. A. Doshier, et al. (1990). "How to study the kinetic depth experimentally " *Journal of Experimental Psychology: Human Perception and Performance* 16: 445-450.
- Stevens, S. S. (1951). *Handbook of Experimental Psychology*. London, Chapman & Hall
- Teller, D. Y. (1984). "Linking Propositions." *Vision Research* 24(10): 1233-1246.

Draft. Please do not quote without permission. Comments welcome.

To, M. P. S., P. G. Lovell, et al. (2008). "Summation of perceptual cues in natural visual scenes." *Proc. Royal. Soc. Lond. B. Biol. Sci* 275: 2299-2308.

To, M. P. S., P. G. Lovell, et al. (2010). "Perception of suprathreshold naturalistic changes in colored natural images." *J. Vision* 10: 1-22.

Tolhurst, D. J., M. P. S. To, et al. (2010). "Magnitude of Perceived Change in Natural Images May be Linearly Proportional to Differences in Neuronal Firing Rates." *Seeing and Perceiving* 23: 349-372. Reprinted in J.A. Soloman (ed) 2011, *Fechner's Legacy in Psychology*. Boston: Brill